**XML and Web Technologies - 201819 [INFO-H509]**


Hypothesis report

# Project 3:XQuery

*Authors:*
Siéwé Kouéta Anicet : 00364245

*Professor:*
Stijn Vansummeren

UNIVERSITÉ LIBRE DE BRUXELLES ULB

May 2, 2019

# Contents

# 1 Introduction

this report aims to give a detailed explanation of the choice of our hypotheses when writing XQuery programs for each of the requests

# 2 First request

The first query is written to the **Queries1.xquery** file. An output from this query is in the **Output_Queries1.xml** file.
At first we have Select all <**author**> nodes of the **dblp-excerpt.xml** file in separate ways.

```
for $SelectAuthor in distinct-values( doc("dblp-excerpt.xml")//*/author )
```

then, Count the number of parent nodes for which the child node **author** = **author_select** appears.

[

```
<coauthors number="{number(count(distinct-values(doc("dblp-excerpt.xml")
//*[author=$SelectAuthor]/author)))-1}">
```

For each author, we display the list of his **co-author**.

[

```
for $Co_Author in distinct-values(doc("dblp-excerpt.xml")//*[author=$SelectAuthor]/author)
```

here for each co-author, we count all the articles in which the **author** appears.

```
{count(distinct-values(doc("dblp-excerpt.xml")//*[author=$SelectAuthor]/author[.=$Co_Author] ))}
```

# 3 Second request

The second query is written to the **Queries2.xquery** file. An output of this query is in the **Output_Queries2.xml** file.

For the output format to suggest in the project statement, we found it necessary to insert an additional node: the node <**Proceedings**> with a capital **'P'**, which will serve as the root element in our **Output_Queries2.xml** output file when the generation of xml code.

We select all the <**proceedings**> nodes of the file **dblp-excerpt.xml**

```
for $SelectProceedings in doc("dblp-excerpt.xml")//proceedings
```

for each parent node <**proceeding**>, we display the child data <**title**>

```
<proc_title>{data($SelectProceedings/title)}</proc_title>
```

For each <**proceeding**> node, we select all <**title**> child nodes of the document that have the child node <**crossref**> with the value of the **'key'** attribute of the <**proceeding**> node in progress.

```
for $Selecttitle in doc("dblp-excerpt.xml")//*[crossref=data($SelectProceedings/@key)]/title
```

# 4 Third request

The third query is written to the **Queries3.xquery** file. An output from this query is in the **Output_Queries3.xml** file.
To calculate all the distance of 1 between two author, we made a double loop.
We select each author, and for each author, we select his co-author. For each co-author to find, we display the author and his co-author with a distance of **1**:

```
for $SelectAuthor in distinct-values( doc("dblp-excerpt.xml")//*/author ),
$Co_Author in distinct-values(doc("dblp-excerpt.xml")//*[author=$SelectAuthor]/author)
```

for each co-author of an author, we select the other co-author of this co-author

```
for $other_Co_Author in distinct-values(doc("dblp-excerpt.xml")//*[author=$Co_Author]/author)
```

and we display the author and the other co-author with a distance of **2**.

to avoid that the other co-author is already a co-author of the author, we have written a function that allows us to check first whether the other co-author is already a co-author of the current author . This function displays the co-author list of an author and checks whether the other co-author makes this list:

- If 'YES', we do not perform any results, because this result is already processed by the first loop.

- If 'NO' returns the display of the author and the other co-author with a distance of **2**.